

Serial No. 09/420,796
Docket No. YO999-421

9

REMARKS

Claims 1-9, 11-14 and 16-25 are all the claims presently pending in the application.

Claims 1-6, 8, 11-13 and 16-25 have been amended to more clearly define the invention. Claims

1, 16, 20 and 23-24 are independent.

Entry of this §1.116 Amendment is proper. Since the amendments above narrow the issues for appeal and since such features and their distinctions over the prior art of record were discussed earlier, such amendments do not raise a new issue requiring a further search and/or consideration by the Examiner. As such, entry of this Amendment is believed proper and Applicant earnestly solicits entry. No new matter has been added.

These amendments are made only to more particularly point out the invention for the Examiner and not for narrowing the scope of the claims or for any reason related to a statutory requirement for patentability.

Applicant also notes that, notwithstanding any claim amendments herein or later during prosecution, Applicant's intent is to encompass equivalents of all claim elements.

Claims 1-9, 11-14 and 16-25 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Kazama, et al. (U.S. Patent No. 6,111,580), in view of Iwamura (U.S. Patent No. 6,498,628).

This rejection is respectfully traversed in the following discussion.

Serial No. 09/420,796
Docket No. YO999-421

10

I. THE CLAIMED INVENTION

The claimed invention is directed to a method of enabling a computer system to recognize specific actions of a user. The method includes capturing an image of a user within a video stream which includes a predetermined target area, displaying the image, determining that a state of the predetermined target area is to be associated with a computer event, associating the computer event with the state of the predetermined target area in response to the determination, and storing information in a memory device regarding the association.

Conventional computers that have been adapted for users having physical disabilities have included devices like a touch screen or a single-switch device. However, these computers have distinct disadvantages. First, they rely upon physical devices that require careful setup for the user. Second, they are prone to damage and/or vandalism. Third, they do not allow the full range of expression needed to effectively interact with computer applications.

Some conventional computer systems are capable of interacting with users by recognizing gestures using cameras. However, these systems are generally very limited in the type of gestures that they are capable of recognizing, they require extensive customization for each user, they are not robust in the face of environmental conditions, they are not reliable or they require extensive user training.

By contrast, the present invention provides a robust, flexible and user friendly method and apparatus which allows a computer to recognize a wide range of user actions using a camera.

Serial No. 09/420,796
Docket No. YO999-421

11

II. THE PRIOR ART REJECTION

Regarding the rejection of claims 1-9, 11-14 and 16-25, the Examiner alleges that the Iwamura reference would have been combined with the Kazama et al. reference to form the claimed invention. Applicants submit, however, that these references would not have been combined and, even if combined, the combination would not teach or suggest each and every element of the claimed invention.

Applicants submit that these references would not have been combined as alleged by the Examiner. Indeed, the references are directed to completely different matters and problems.

Specifically, the Kazama et al. reference is directed to preventing inadvertent operation of a computer device which includes a gesture and/or voice recognition input device by providing a system that determines whether a user's attention is actually being directed toward the computer device. The Kazama et al. reference discloses a system that first determines whether a user is actively operating the system and, if the system determines that the user is actively operating the system, determines whether the user performs an act that corresponds to a command. The Kazama et al. reference discloses a gaze direction detector which determines whether a user is actively operating the system based upon whether an image of a user indicates that the user is directing attention to the system (col. 3, line 16 - col. 5, line 25). The Kazama et al. reference also appears to disclose determining the attention of a user based upon an analysis of the user's voice (col. 6, lines 32 - 44) or through direct contact (col. 6, lines 45 - 59). The Kazama et al. reference also appears to disclose a gesture recognition section that requires the user to wear a

Serial No. 09/420,796
Docket No. YO999-421

12

glove including colored balls at the end of each finger of the user (col. 5, lines 26 - 65 and col. 9, lines 21 - 42).

In contrast, the Iwamura reference is specifically directed to solving the problems of operating electronic appliances using remote controllers by providing a system which presents icons representing possible operations of an electronic appliance and a motion detector circuit which detects either an image of a user's hand or the motion of a user's moving hand (col. 1, lines 35-45). Therefore, since the Kazama et al. reference is directed to preventing inadvertent operation of an electronic device rather than to being directed to obviating the problems of remote controllers as disclosed by the Iwamura reference, one of ordinary skill in the art would not have been motivated to modify the system disclosed by the Kazama et al. reference in view of the system disclosed by the Iwamura reference because they are directed to entirely different problems and matters. Thus, the references would not have been combined, absent hindsight.

Further, Applicant submits that the Examiner can point to no motivation or suggestion in the references to urge the combination as alleged by the Examiner. Indeed, the Examiner does not even support the combination by identifying a reason for combining the references.

Rather, the Examiner merely alleges that it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system disclosed in the Kazama et al. reference with the active regions disclosed by the Iwamura reference "because it allows a user to input an operation without use of any operating section for inputting data other than position data." There are at least three problems with the Examiner's alleged motivation.

Serial No. 09/420,796
Docket No. YO999-421

13

Firstly, the Examiner has failed to provide a citation to any of the applied references which provides a basis for the alleged motivation. Indeed, Applicant respectfully submits that such a motivation may only be found in the Applicant's own disclosure. Clearly, the Examiner has engaged in the impermissible use of hindsight in an attempt to provide a motivation for the alleged modification.

Secondly, the alleged modification of providing an active region to the system disclosed by the Kazama et al. reference does not "allow a user to input an operation without the use of any operation section for inputting data other than position data." The Iwamura reference does not teach or suggest that the determination and use of "active areas" is what provides the alleged benefit.

Thirdly, the alleged motivation is completely inapplicable to the system disclosed by the Kazama et al. reference. Indeed, the Kazama et al. reference is fully capable of providing the benefit alleged by the Examiner without modification. Therefore, there is no reason at all to modify the system disclosed by the Kazama et al. reference to "allow a user to input an operation without the use of any operation section for inputting data other than position data" because the system disclosed by the Kazama et al. reference already provides this benefit.

Even assuming arguendo that one of ordinary skill in the art would have been motivated to combine these references, the combination would not teach or suggest each and every element of the claimed invention. The Examiner admits that the Kazama et al. reference does not teach or suggest a predetermined target area.

Serial No. 09/420,796
Docket No. YO999-421

14

The Iwamura reference, like the Kazama et al. reference, also does not teach or a predetermined target area. The claimed invention provides a predetermined target area in which a user image is captured, displayed and used to associate with a computer event. For example, an exemplary embodiment of the invention provides a predetermined target area which is static (i.e. stationary). The use of a predetermined target area avoids the need to track and interpret activity throughout the image being received which significantly reduces the processing requirements and also the risk of false positive events.

The present invention provides a predetermined target area. In this manner, only the predetermined target area requires processing and then only the activities which are relevant are detected. This allows for the creation of interfaces which are built up as collections of targets that the user triggers as needed. These predetermined target areas can be placed in regions of the image that correspond to visible objects in the environment so that the user can trigger the predetermined target by, for example, touching the visible object within the predetermined target area.

Alternatively, the predetermined target areas may correspond to regions in the image near the user's current position so that the users motion in a particular direction within the predetermined target area will trigger the computer event.

By contrast, the Iwamura reference discloses an "active area" which is a region of the image where the system detects movement which may be a moving hand. These active areas are detected using "macro blocks", (a concept from MPEG video compression, which are simply, static 16 x 16 pixel blocks that tile the video image). The Iwamura reference discloses

Serial No. 09/420,796
Docket No. YO999-421

15

examining each macro block for a change between frames. If a sufficient change is detected in a macro block, it is labeled as being active. Spatially adjacent active macro blocks are collected into active areas. Active areas are then examined to determine if they are possible hand candidates, generally by looking for the pattern of movement between video frames. Thus, the "active areas" disclosed by the Iwamura reference are moving regions of the image where the system has detected movement in the recent past.

The Office Action refers to Figure 4 of the Iwamura reference. This figure illustrates an active area composed of a rectangle of four macro blocks which have detected a change in a video frame and where the active area moves over time. Clearly, this does not teach or suggest a predetermined target area.

The Office Action further refers to Figures 6-8 of the Iwamura reference. These figures illustrate how a hand movement can generate a corresponding movement in an active area composed of clusters of active macro blocks. Again, this feature has absolutely nothing to do with a predetermined target area.

Furthermore, the Office Action refers to Figure 9 of the Iwamura reference. This figure illustrates hand motion (which generate moving active areas in the processed video stream) being used to move a cursor on a TV screen and buttons within that TV screen which wait for the cursor to trigger them. Again, these buttons have no relation at all to the claimed predetermined target area. The predetermined target areas of the present invention are predetermined. By contrast, the buttons of Figure 9 are static regions on the TV image stream which is completely different.

Serial No. 09/420,796
Docket No. YO999-421

16

Lastly, Applicants believe that a brief review of the technology may be of assistance. The goal here is to use a camera to identify events such as a user's hand gestures or body movements and to generate events within a computer once those events are identified. The computer events may be the movement of a mouse on a screen or trigger some computer activity.

There are three frames of reference which should be considered: 1) the real World, where the user performs the actions; 2) the image stream of a Camera observing the user in the real world; and 3) the image stream for a Screen, where the user can observe the activity of the vision system and the results of the generated computer event. These will be referred to as "World," "Camera" and "Screen," respectively.

There does not appear to be any misunderstanding of the World. Therefore, Applicants will not address that frame of reference. However, there does appear to be a misunderstanding between the Camera and Screen frames of reference.

It is important to note that an exemplary embodiment of the present invention examines activity in the Camera image stream and can generate activity in the Screen image stream, in response, but that the two frames of reference are not related in any other way.

Figure 9 of the Iwamura reference illustrates how a user motion in the World is recognized by tracking that motion with the active area in the Camera image stream which, in turn, is used to move a pointer on the Screen image stream. The user observes the pointer and moves their hand in such a way as to move the pointer into a button on the Screen image stream and so generate some event. While both the claimed target area and the button displayed in the Screen image stream are regions where activity can generate a computer event, note that the

Serial No. 09/420,796
Docket No. YO999-421

17

predetermined target area of an exemplary embodiment of the invention identifies an area in the Camera image stream where observed events must occur. By contrast, the button disclosed in the Iwamura reference identifies an area in the Screen image stream where computer events must occur. Thus, the two are completely different entities.

A second misunderstanding associated with these frames of reference is that the Camera image stream may be displayed in the Screen image stream. When this occurs, the target area may be drawn on top of the Camera image stream, but that should not be taken to require that it represents a region of the Screen image stream because if the location where the Camera image stream is drawn is moved, the target area moves with it. In an exemplary embodiment of the invention, the target area is always located with respect to the Camera image stream, regardless of where or if the target area is displayed in the Screen image stream.

In summary, the present invention provides predetermined target areas which serve to localize the region in which a system examines for specific image events. By contrast, the active areas disclosed by the Iwamura reference are moving regions of the image where the system believes there may be a hand. Clearly, the Iwamura reference does not teach or suggest a predetermined target area.

Clearly, this novel feature are not taught or suggested by the applied references. Indeed, the applied references is completely unrelated to the claimed invention.

Therefore, the Examiner is respectfully requested to withdraw this rejection of claims 1-9, 11-14 and 16-25.

Serial No. 09/420,796
Docket No. YO999-421

18

III. FORMAL MATTERS AND CONCLUSION

In view of the foregoing amendments and remarks, Applicant respectfully submits that claims 1-9, 11-14 and 16-25, all the claims presently pending in the Application, are patentably distinct over the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the Application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

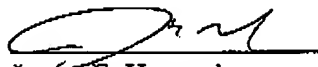
Serial No. 09/420,796
Docket No. YO999-421

19

The Commissioner is hereby authorized to charge any deficiency in fees or to credit any overpayment in fees to Attorney's Deposit Account No. 50-0481.

Respectfully Submitted,

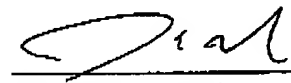
Date: 5/5/03


James E. Howard
Registration No. 39,715

McGinn & Gibb, PLLC
8321 Old Courthouse Rd., Suite 200
Vienna, Virginia 22182
(703) 761-4100
Customer No. 21254

CERTIFICATION OF FACSIMILE TRANSMISSION

I hereby certify that I am filing this Amendment by facsimile with the United States Patent and Trademark Office to Examiner Crescelle N. Dela Torre, Group Art Unit 2174 at fax number (703) 746-7238 this 5th day of May, 2003.


James E. Howard
Reg. No. 39,715